

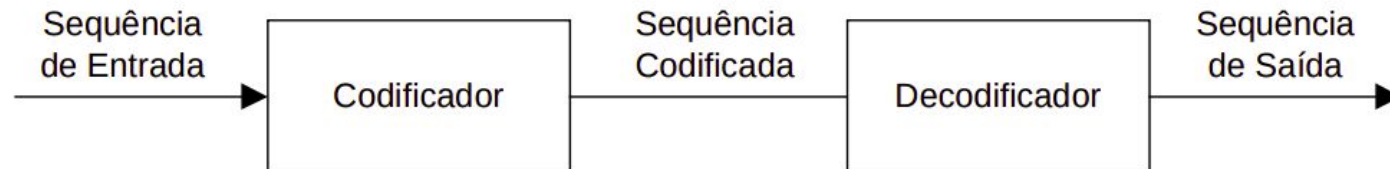


Compressão LZW

Prof. Jose J. Camata

Introdução

- Modificação do algoritmo LZ proposto por Lempel e Ziv (1977)
- Ideia principal é ir construindo um dicionário de símbolos ou palavras conforme o texto ou a informação vai sendo processado pelo algoritmo
- Possui uma tabela como dicionário que é inicializada com todos os símbolos da fonte
 - Por exemplo, pode-se inicializar o dicionário com os caracteres da tabela ASCII
- O dicionário não tem limite de entradas
- O código gerado é um único valor que indica o índice no dicionário





Algoritmos

As convenções adotadas são:

P	string que representa um prefixo
C	caracter
$X + Y$	Concatena a string X com Y
SC	Saída Codificada
cW	Código de uma palavra
PW	Código de um prefixo
string(cW)	string associada ao código cW
codigo(S)	código associado a string S

Codificação

1. No início o dicionário contém todas as símbolos possíveis
2. $P \leftarrow \emptyset$, $SC \leftarrow \emptyset$
3. $C \leftarrow$ recebe o primeiro caractere da entrada
4. A string $P + C$ existe no dicionário?
 - 4.1. Se verdade
 - 4.1.1. Atualize o prefixo: $P \leftarrow P + C$
 - 4.2. Se falso,
 - 4.2.1. Escreva o código de P na saída: $SC \leftarrow SC + \text{código}(P)$
 - 4.2.2. Adicione a string $P+C$ ao dicionário;
 - 4.2.3. Atualize o prefixo: $P \leftarrow P + C$
5. Existem mais caracteres na sequência de entrada ?
 - 5.1. Se verdade, $C \leftarrow$ caractere da entrada, e volte para o passo 4
 - 5.2. Se falso, escreva o código de P na saída: $SC \leftarrow SC + \text{código}(P)$ e SAIA;

Exemplo

Deseja-se codificar a sequência abaixo:

Inicializamos o dicionário

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC								

1. $SC \leftarrow \emptyset$
2. $P \leftarrow \emptyset$
3. $C \leftarrow a$
4. A string $P + C$ ("a") existe no dicionário?
 - a. $P \leftarrow \emptyset + a$
5. Existem mais caracteres na sequência de entrada ?
 - a. Sim, $C \leftarrow b$ e vai para passo 4

Código	Símbolo
1	a
2	b
3	c

Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	1							

Código	Símbolo
1	a
2	b
3	c
4	ab

$P \leftarrow a$

$C \leftarrow b$

4. A string $P + C$ ("ab") existe no dicionário? Não
 - a. Escreva código (P) em SC
 - b. Adicione ab no dicionário
 - c. $P \leftarrow C$
5. Existem mais caracteres na sequência de entrada ?
 - a. Sim, $C \leftarrow b$ e vai para passo 4

Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	12							

$P \leftarrow b$

$C \leftarrow b$

4. A string $P + C$ ("bb") existe no dicionário? Não
 - a. Escreva código (P) em SC
 - b. Adicione bb no dicionário
 - c. $P \leftarrow b$
5. Existem mais caracteres na sequência de entrada ?
 - a. Sim, $C \leftarrow a$ e vai para passo 4

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb

Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	122							

$P \leftarrow b$

$C \leftarrow a$

4. A string $P + C$ ("ba") existe no dicionário? Não
 - a. Escreva código (P) em SC
 - b. Adicione ba no dicionário
 - c. $P \leftarrow a$
5. Existem mais caracteres na sequência de entrada ?
 - a. Sim, $C \leftarrow b$ e vai para passo 4

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba

Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	122							

$P \leftarrow a$

$C \leftarrow b$

4. A string $P + C$ ("ab") existe no dicionário? Sim
 - a. $P \leftarrow a + b$
5. Existem mais caracteres na sequência de entrada ?
 - a. Sim, $C \leftarrow a$ e vai para passo 4

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba

Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	1224							

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba
7	aba

$P \leftarrow ab$

$C \leftarrow a$

4. A string $P + C$ ("aba") existe no dicionário? não
 - a. Escreva código (P) em SC
 - b. Adicione aba no dicionário
 - c. $P \leftarrow a$
5. Existem mais caracteres na sequência de entrada ?
 - a. Sim, $C \leftarrow b$ e vai para passo 4

Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	1224							

$P \leftarrow a$

$C \leftarrow b$

4. A string $P + C$ ("ab") existe no dicionário? sim
 - a. $P \leftarrow a + b$
5. Existem mais caracteres na sequência de entrada ?
 - a. Sim, $C \leftarrow a$ e vai para passo 4

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba
7	aba

Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	1224							

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba
7	aba

$P \leftarrow ab$

$C \leftarrow a$

4. A string $P + C$ ("aba") existe no dicionário? sim
 - a. $P \leftarrow ab + a$
5. Existem mais caracteres na sequência de entrada ?
 - a. Sim, $C \leftarrow c$ e vai para passo 4



Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	12247							

$P \leftarrow \text{aba}$

$C \leftarrow \text{c}$

4. A string $P + C$ ("abac") existe no dicionário? não
 - a. Escreva código (P) em SC
 - b. Adicione aba no dicionário
 - c. $P \leftarrow c$
5. Existem mais caracteres na sequência de entrada ?

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba
7	aba
8	abac



Exemplo

Deseja-se codificar a sequência abaixo:

0	1	2	3	4	5	6	7	8
a	b	b	a	b	a	b	a	c
SC	122473							

$P \leftarrow \text{aba}$

$C \leftarrow \text{c}$

4. A string $P + C$ (“abac”) existe no dicionário? não
 - a. Escreva código (P) em SC
 - b. Adicione aba no dicionário
 - c. $P \leftarrow c$
5. Existem mais caracteres na sequência de entrada ?
 - a. não, Escreva código (P) em SC e FIM

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba
7	aba

Decodificação

1. No início o dicionário contém todas as símbolos possíveis
2. $cW \leftarrow$ primeiro código de SC
3. Coloque $\text{string}(cW)$ na sequência de saída.
4. $pW \leftarrow cW$
5. $cW \leftarrow$ próximo código de SC
6. A $\text{String}(cW)$ existe no dicionário?
 - a. Se sim,
 - i. Coloque a $\text{string}(cW)$ na sequência de saída
 - ii. $P \leftarrow \text{string}(pW)$
 - iii. $C \leftarrow$ primeiro caractere da $\text{string}(cW)$
 - iv. Adicione a string $P + C$ ao dicionário
 - b. Se não,
 - i. $P \leftarrow \text{string}(pW)$
 - ii. $C \leftarrow$ primeiro caractere da $\text{string}(cW)$
 - iii. Adicione a string $P + C$ ao dicionário e escreva na saída.
7. Existe mais códigos em SC?
 - a. se sim, volte ao passo 4
 - b. se não, FIM

Exemplo

Deseja-se decodificar a sequência abaixo:

SC	122473
SS	a b

Código	Símbolo
1	a
2	b
3	c
4	ab

SS $\leftarrow \emptyset$

cW $\leftarrow 1$

Coloque string(cW) em SS

pW Coloque string(cW) na sequência de saída $\leftarrow 1$

cw $\leftarrow 2$ (próximo código de SC)

A string(cW) = "b" existe no dicionário? sim

Coloque string(cW) em SS

P $\leftarrow a$

C $\leftarrow b$

adicione a+b no dicionário

Existe mais códigos em SS? sim, volte para passo 4 do algoritmo



Exemplo

Deseja-se decodificar a sequência abaixo:

SC	122473
SS	abb

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb

$pW \leftarrow 2$

$cW \leftarrow 2$ (próximo código de SC)

A string(cW) = "b" existe no dicionário?

Coloque string(cW) em SS

$P \leftarrow b$

$C \leftarrow b$

adicione $b+b$ no dicionário

Existe mais códigos em SS? sim.

volte para passo 4 do algoritmo

Exemplo

Deseja-se decodificar a sequência abaixo:

SC	122473
SS	abbab

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba

$pW \leftarrow 2$

$cW \leftarrow 4$ (próximo código de SC)

A string(cW) = "ab" existe no dicionário? sim

Coloque string(cW) em SS

$P \leftarrow \text{string}(pW) = "b"$

$C \leftarrow \text{primeira letra}(\text{string}(cW)) = "a"$

adicione b+a no dicionário

Existe mais códigos em SS? sim.

volte para passo 4 do algoritmo

Exemplo

Deseja-se decodificar a sequência abaixo:

SC	122473
SS	abbababa

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba
7	aba

$pW \leftarrow 4$

$cW \leftarrow 7$ (próximo código de SC)

A string(cW) = "ab" existe no dicionário? não

$P \leftarrow \text{string}(pW) = \text{"ab"}$

$C \leftarrow \text{primeira letra}(\text{string}(pW)) = \text{"a"}$

Coloque string($P+C$) em SS

adicione ab+ a no dicionário

Existe mais códigos em SS? sim.

volte para passo 4 do algoritmo

Exemplo

Deseja-se decodificar a sequência abaixo:

SC	122473
SS	abbababac

Código	Símbolo
1	a
2	b
3	c
4	ab
5	bb
6	ba
7	aba
8	abac

$pW \leftarrow 7$

$cW \leftarrow 3$ (próximo código de SC)

A string(cW) = "c" existe no dicionário? sim

Coloque string(cW) = "c" em SS

$P \leftarrow (\text{string}(pW)) = \text{"aba"}$

$C \leftarrow \text{primeira letra}(\text{string}(cW)) = \text{"c"}$

adicione $\text{aba}+c$ no dicionário

Existe mais códigos em SS? não.

FIM

Exercício:

Aplique os algoritmos de codificação e decodificação na seguinte cadeia de caracteres: LAILAALAALAALAA